

| | | | | | | | | | | | | | | | | | | | |
|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|
| | | | | | | | | | | | | | | | | | | | |
|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|

NOIDA INSTITUTE OF ENGINEERING AND TECHNOLOGY, GREATER NOIDA

(An Autonomous Institute Affiliated to AKTU, Lucknow)

B.Tech

SEM:VI CARRY OVER THEORY EXAMINATION-AUGUST 2023

Subject: Programming for Data Analytics

Time: 3 Hours

Max. Marks: 100

General Instructions:

IMP: Verify that you have received the question paper with the correct course, code, branch etc.

1. This Question paper comprises of **three Sections -A, B, & C.** It consists of Multiple Choice Questions (MCQ's) & Subjective type questions.
2. Maximum marks for each question are indicated on right -hand side of each question.
3. Illustrate your answers with neat sketches wherever necessary.
4. Assume suitable data if necessary.
5. Preferably, write the answers in sequential order.
6. No sheet should be left blank. Any written material after a blank sheet will not be evaluated/checked.

SECTION A

20

1. Attempt all parts:-

- 1-a. How do you access a specific element in a Series in Pandas?(CO1) 1
- (a) Using .iloc[] accessor function
 - (b) Using indexing operator [] and passing in the name of the column
 - (c) Using .loc[] accessor function
 - (d) Using indexing operator [] and passing in the label or integer location of the element
- 1-b. Which of the following is not a method to read data into a DataFrame in Pandas?(CO1) 1
- (a) pd.read_csv()
 - (b) pd.read_excel()
 - (c) pd.read_table()
 - (d) pd.read_html()
- 1-c. An .Rproj file is a ___ file.(CO2) 1
- (a) pdf

- (b) text
- (c) jpeg
- (d) png
- 1-d. _____ is simplest class of analytics.(CO2) 1
- (a) Predictive
- (b) Descriptive
- (c) Prescriptive
- (d) Summarization
- 1-e. With SQLite, there is no ____ required, making it an excellent choice for small businesses.(CO3) 1
- (a) Configuration
- (b) Administration
- (c) Both A and B
- (d) None of the above
- 1-f. MongoDB is a _____ database that provides high performance, high availability, and easy scalability.(CO3) 1
- (a) graph
- (b) key value
- (c) document
- (d) all of the mentioned
- 1-g. Which TensorFlow API allows for distributed training across multiple devices?(CO4) 1
- (a) TensorFlow Lite
- (b) TensorFlow.js
- (c) TensorFlow Estimators
- (d) TensorFlow Distributed
- 1-h. What is a queue in TensorFlow?(CO4) 1
- (a) A data structure used for storing tensors
- (b) A tool for visualizing TensorFlow graphs
- (c) A type of machine learning algorithm
- (d) None of the above
- 1-i. Which of the following is a type of layer commonly used in convolutional neural networks?(CO5) 1

- (a) Linear Layer
- (b) Pooling Layer
- (c) Fully connected layer
- (d) Parallel Layer

- 1-j. Which of the following is a commonly used application of deep reinforcement learning?(CO5) 1
- (a) Autonomous driving
 - (b) Image generation
 - (c) Speech recognition
 - (d) Voice recording

2. Attempt all parts:-

- 2.a. How can regular expressions be used for data validation?(CO1) 2
- 2.b. What are the data structures in R that is used to perform statistical analyses and create graphs? (CO2) 2
- 2.c. Why do we use the pretty() method?(CO3) 2
- 2.d. What is a TensorFlow session?(CO4) 2
- 2.e. What is the encoder in an autoencoder?(CO5) 2

SECTION B

30

3. Answer any five of the following:-

- 3-a. Write a Python program to create a Pandas DataFrame from a dictionary of lists, where the keys are the column names and the values are the data for each column.(CO1) 6
- 3-b. Write a Python program to select only the rows from a DataFrame where the 'gender' column is 'Male' and the 'age' column is greater than 30.(CO1) 6
- 3-c. What does geom_col() do? How is it different to geom_bar()?(CO2) 6
- 3-d. What the following tools do in regular expression of R?(CO2) 6
- a) \d
 - b) \s
 - c) [abc]
 - d) [^abc]
- 3.e. Explain Namespace in MongoDB?(CO3) 6
- 3.f. What is reinforcement learning, and how is it implemented in TensorFlow?(CO4) 6
- 3.g. What is the architecture of a GAN?(CO5) 6

4. Answer any one of the following:-

- 4-a. Write python program to filter out all email address form a text file using regular expression.(CO1) 10
- 4-b. What is the difference between a one-tailed and a two-tailed p-test? (CO1) 10

5. Answer any one of the following:-

- 5-a. What is List in R? Explain creation, accessing and manipulation of List with example in R.(CO2) 10
- 5-b. What are the steps to build and evaluate a linear regression model in R?(CO2) 10

6. Answer any one of the following:-

- 6-a. Explain MongoDB Shell commands with syntax and example.(CO3) 10
- 6-b. Explain Comparison Query Operator with example in MongoDB.(CO3) 10

7. Answer any one of the following:-

- 7-a. What are some common pre-processing techniques used for text and sequence data?(CO4) 10
- 7-b. Explain the working of Convolutional Neural Network?(CO4) 10

8. Answer any one of the following:-

- 8-a. What are some common challenges associated with GANs, such as mode collapse, instability, and hyperparameter tuning, and how can they be mitigated?(CO5) 10
- 8-b. What is Deep Reinforcement Learning, and how does it differ from traditional Reinforcement Learning?(CO5) 10